

***STRUCTURE COMMUNICATIVE ET GEOMETRIE
INTONATIVE : QUE NOUS DIT LA SYNTHÈSE DE LA
PAROLE ?***

Anne Lacheret-Dujour
Université de Caen, France

ABSTRACT

A l'heure où la génération automatique de la prosodie repose encore uniquement sur la prise en compte de contraintes syntaxiques et rythmiques, cela pour des résultats relativement médiocres, il faut avouer, il apparaît nécessaire de s'interroger sur l'apport d'informations pragmatiques pour améliorer le module de génération. Cette axe de réflexion conduit inévitablement à définir aussi précisément que possible les relations entre prosodie, sémantique et pragmatique d'une part, prosodie et cognition d'autre part. En aval, il révèle de nouveaux paradigmes de recherche concernant les corrélats acoustiques objectifs de ces différentes composantes. La synthèse de la parole trouve alors toute sa justification scientifique en tant qu'outil de simulation utilisé pour vérifier un certain niveau d'adéquation entre hypothèses théoriques et réalisations concrètes.

1. INTRODUCTION

La thématique générale de notre recherche concerne les niveaux d'analyse et les problèmes de représentation formelle de la prosodie des langues dans une perspective d'interprétation fonctionnelle et cognitive. Sous l'angle fonctionnel

d'abord, il s'agit de rendre compte aussi rigoureusement que possible des relations entre intonation, syntaxe, sémantique et pragmatique. Du point de vue cognitif ensuite, nous souhaitons préciser les processus mentaux qui sous-tendent les configurations intonatives rencontrées en parole et la manière dont celles-ci sont traitées (caractère intégré ou autonome du module prosodique par rapport aux autres modules de traitement). En pratique, rendre compte des constructions prosodiques suppose de développer des dispositifs formels pertinents en production comme en réception, qui rendent compte du rapport entre états mentaux et productions sonores attestées¹. Nous travaillons donc à la fois du côté de la génération et de l'analyse ou, plutôt, de l'analyse et de la génération, puisque l'observation des patrons prosodiques en français parlé nous conduit à émettre un ensemble d'hypothèses sur leurs conditionnements linguistiques et cognitifs, et à utiliser ensuite la synthèse de la parole dans des tâches de perception, de manière à tester leur pertinence. Ces hypothèses, quelles sont-elles et dans quelle mesure peuvent-elles être formalisées par des règles de génération automatique ? Tel est l'objet de notre présentation, que nous illustrerons ici par l'étude du corpus Fayard-Groult.

2. HYPOTHESES, MATERIEL DEFINITOIRE ET METHODE DE TRAVAIL

La remise en cause du primat et de l'autonomie de la syntaxe pour la dérivation des structures prosodiques constitue le point d'ancrage de nos travaux. De ce point de vue, nous nous écartons significativement des modèles intonologiques classiques qui établissent toujours plus ou moins explicitement un lien entre les structures intonative et syntaxique, et effectuent une segmentation prosodique sur les bases de critères syntaxiques aménagés, le cas échéant, pour tenir compte de la dimension rythmique du message. C'est également cette approche qui est utilisée pour la génération automatique de la prosodie dans les synthétiseurs de parole. Est-ce à dire qu'un tel traitement dérive d'une réalité cognitive ? Bien évidemment non : pour les constructeurs des systèmes de synthèse, la syntaxe n'est, en quelque sorte, que la poire pour la soif, utilisée dans l'optique de générer une prosodie « par défaut » ou « à défaut » en attendant mieux. Mais qu'est-ce que ce mieux ? Est-il raisonnable d'attendre ? Le « mieux » est-il modélisable et implémentable dans un système de génération automatique de la prosodie ? Autrement dit, comment passer de descriptions fines, très fines, sur la matière énonciative (Culioli 1990, Morel & Danon-Boileau 1998, Berrendonner 2002) à des réalisations objectives, puisqu'en définitive

¹ Cet angle d'attaque s'inscrit directement dans le cadre des recherches en phonologie cognitive (Voir Laks 1997 pour une discussion).

c'est bien de cela qu'il s'agit ? N'y a-t-il pas des aspects qui résistent à l'implémentation et qui, par nature, resteront toujours de l'ordre de l'indicible pour la machine ? Bref, comment passer de contraintes pragmatiques, sémantiques et topologiques à un balisage textuel pertinent d'une part, dans quelle mesure un tel traitement est-il implémentable et donc exploitable pour une tâche de génération automatique de la prosodie d'autre part ?

2.1. Enjeux sémantiques et pragmatiques

Le parti-pris de renoncer à un éclairage syntaxique pour rendre compte du formatage des objets prosodiques en parole dérive directement d'un fait empirique incontournable : l'observation des constructions intonatives révèle de nombreux contextes de violation syntaxique. Ainsi dans :

j'étais d'une famille très unie très aisée

la segmentation intonative :

(j'étais d'une famille très unie) (très aisée)

ne correspond pas à l'organisation en constituants

(j'étais) ((d'une famille) (très unie) (très aisée)).

Nous posons donc comme **première hypothèse** que la production des structures prosodiques ne suppose pas un encodage syntaxique préalable et proposons un traitement intégré de la syntaxe et de la prosodie. Autrement dit, il s'agit là de deux structures qui coopèrent, le cas échéant par le biais de rendez-vous structurels, pour la construction du sens. Elles se configurent donc fondamentalement en réponse aux instructions provenant des niveaux supérieurs de traitement : les niveaux symbolique et communicatif² (figure 1).

² Ce point justifie la notion de **identité dynamique** développée en *infra* à la fin de cette section.

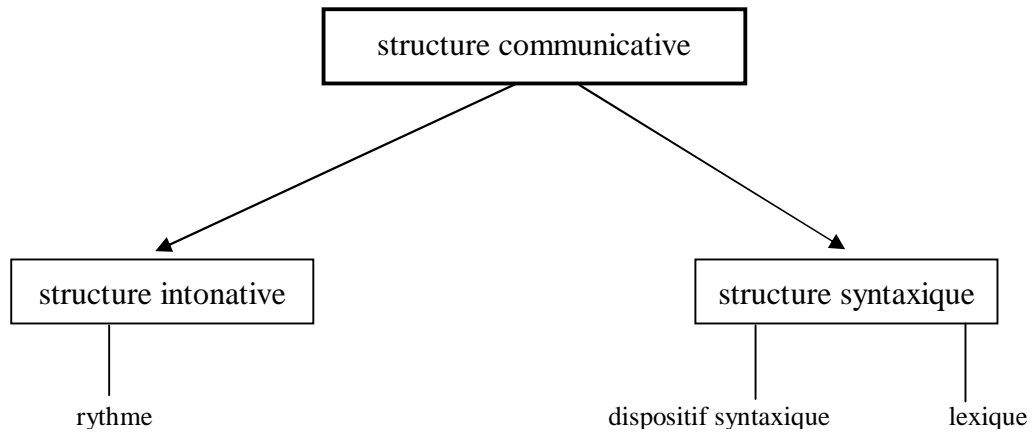


Figure 1 : l'intonation et la syntaxe, des systèmes adaptatifs

où : outre les instructions projetées par la structure communicative sur la structuration intonosyntaxique du message, les contraintes internes à chaque niveau de traitement sont : le rythme pour la structure intonative, les règles de linéarisation et le lexique pour la syntaxe.

Concernant l'aspect syntaxique, cette approche n'est pas nouvelle : elle constitue une des caractéristiques essentielles des grammaires fonctionnelles. Soulignons à cet égard, la théorie de van Valin et Lapolla (1997) pour qui la structure en constituants dérive de trois niveaux de projection : la projection fonctionnelle, codant différents opérateurs sémantiques réalisés par les affixes, auxiliaires et adverbes (aspect, modalité, négation, temps), la structure logique, où sont codées notamment les différentes classes de prédicats verbaux, et la structure pragmatique qui organise l'articulation «topic-comment» telle qu'elle est envisagée par Lambrecht (1994). Quant à la prosodie, si cet angle d'attaque a également été discuté et justifié chez les phonologues (Selkirk 2001), l'autonomie de l'intonation par rapport à la syntaxe constitue bien souvent encore une déclaration de principe, les critères morphosyntaxiques restant toujours le point de départ des architectures prosodiques (pour une présentation, voir Lacheret & Beaugendre 1999 : 88, 89). C'est ce que nous cherchons à éviter soigneusement dans notre entreprise. De ce point de vue, nous posons comme **deuxième hypothèse** que le sens d'un énoncé ne résulte pas d'une vérité pré-existante mais implique une schématisation, co-construite par les sujets (Grize 1996). Cette dernière s'exprime dans les différents niveaux de structuration du message dont, bien sûr, le formatage des objets prosodiques ; c'est ainsi que nous parlons de **schématisation intonative**. A cette hypothèse sont associés deux concepts centraux, le concept de **scène verbale** et celui de **saillance discursive**, sur lesquels il nous paraît nécessaire de nous arrêter un instant.

Voir l'acte d'énonciation comme un processus de construction de scène, c'est dire que :

Au cours de l'activité de parole, se produit un processus cognitif très particulier, qui consiste pour le locuteur à essayer de construire un espace extérieur à lui-même, partageable par ses interlocuteurs, et dans lequel il donne à voir ce dont il parle comme une *scène* projetée devant eux. L'acte d'énonciation de base serait donc constitué d'au moins deux éléments : la *description d'une scène* et la *donnée d'un point de vue sur cette scène*, ce deuxième élément étant essentiel pour permettre le partage de la vision. Victorri & Fuchs (1996, 200).

La notion de scène verbale explique donc celle de saillance discursive : l'énoncé se construit autour de contrastes syntagmatiques utilisés pour jouer sur la saillance des éléments et distinguer, avec des degrés variables, des formes fortes ou **figures** ou à des formes faibles ou **fonds** ou bref, d'opposer premiers plans et arrières plans. Tel est le pouvoir de la prosodie : échapper au carcan de la linéarité, créer perspective et profondeur. Notre question sera donc la suivante : quels sont les indices acoustiques de ces contrastes et quelles sont les ressources phonétiques mobilisées pour créer des îlots de saillance perceptive ? Mais avant de formuler une réponse, encore faut-il définir aussi rigoureusement que possible le concept de **structure communicative** étroitement associé à celui de **schématisation intonative**. Nous la voyons comme une structure qui s'organise autour de deux sous-composantes : (i) une **structure informationnelle** et (ii) une **structure discursive** orthogonale à la première (figure 2). En pratique, pour traiter de la structure informationnelle, nous nous fondons, nous aussi, sur la théorie de Lambrecht selon qui un énoncé est pragmatiquement structuré autour de deux concepts centraux : un élément donné dont on dit quelque chose, accessible par le cotexte discursif ou le contexte situationnel, que l'auteur appelle *topic* et une information nouvelle qu'il dénomme *focus*. Une précision toutefois : il s'agit là d'une conception relationnelle, un élément apportant une information non pas parce qu'il est nouveau mais parce que la relation entre cet élément et le contexte est nouvelle (Lambrecht 1994). En d'autres termes, ce n'est pas le caractère connu d'une entité qui conditionne la notion d'information. Soit l'énoncé *c'est Charlotte qui a mangé les bonbons*, *Charlotte* est un référent connu, néanmoins sa relation au prédicat verbal constitue une information nouvelle. L'élément connu ici relève de la présupposition (*quelqu'un a mangé les bonbons*), l'information nouvelle prenant la forme d'une proposition abstraite ou assertion ou ici : *le X qui a mangé les bonbons est Charlotte*. A la présupposition déjà connue, ou supposée comme telle par le locuteur, s'oppose l'assertion, c'est-à-dire la partie de l'énoncé qui ne sera connue de l'interlocuteur qu'au terme de l'émission du message. Concernant la structure discursive ensuite, elle regroupe essentiellement les marqueurs d'enchaînement énonciatif, les stratégies de négociation dialogale, les indices de point de vue, enfin les phénomènes de coloration affective essentiellement manifestés par l'emphase prosodique.

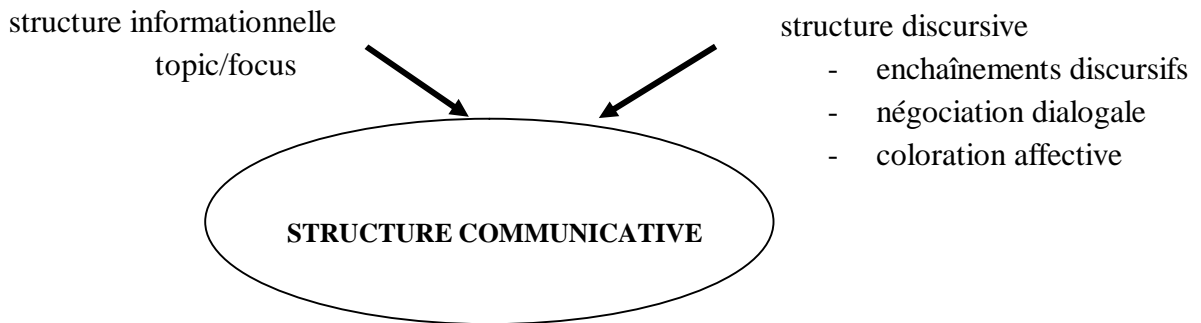


Figure 2 : les composantes de la structure communicative

Notre seconde hypothèse ainsi posée explique l'exemple fourni en *supra*, ((j'étais d'une famille très unie) (très aisée)) et un nouveau positionnement par rapport aux structures : les patrons prosodiques ne peuvent plus être considérés comme des éléments stables, déterminés syntaxiquement une bonne fois pour toute. Ils sont tout au contraire vus comme des entités dynamiques qui ont leur part d'indétermination et se structurent en fonction des relations conceptuelles dans lesquelles ils entrent, et conformément aux instructions projetées par la structure communicative. Les schémas intonatifs ont donc un potentiel de déformabilité dont il faut pouvoir rendre compte. Ainsi, dans notre exemple, le détachement prosodique du second modifieur adjectival s'explique par la saturation sémantique de la tête nominale qui permet au locuteur d'isoler ce constituant. Son autonomie accentuelle, associée à un contour terminal dans le niveau suraigu correspondent à des configurations optimales pour l'instanciation d'un focus étroit (cf. *infra*, § 2.4.) au sens où cette notion est définie par Ladd (1991)³.

2.2. Modélisation topologique des structures

Lorsque l'on travaille sur du matériau textuel, oral comme écrit, nous sommes inévitablement confrontés à une chaîne linéaire séquentielle qui ne peut être ramenée à une simple projection de la structure :

³ Le concept de focus étroit n'est pas sans rappeler celui de **fenêtrage d'attention** proposé par Talmy (2000). Pour rappel, un focus (informationnel) est large lorsqu'il concerne un énoncé complet ou une suite de syntagmes, il est étroit s'il sert à démarquer un seul syntagme, voire un seul item lexical. Plus un focus est étroit, plus il sera interprété comme la marque d'un contraste paradigmatique.

Dans l'énonciation, le locuteur doit projeter sur un axe *linéaire* une pensée *multidimensionnelle* et la *discrétiser* en unités séquentielles. Du fait des propriétés physiques du langage qui est un matériau sonore produit séquentiellement dans le temps, la verbalisation suppose donc de faire passer la pensée par un code particulier qui constitue un goulet d'étranglement (Robert 1997 : 30).

Or, il faut bien faire la part entre les marqueurs associés à l'axe de déroulement syntagmatique et ceux qui, selon la formule de Blanche-Bevéniste (1991), portent les traces d'un piétinement paradigmatique. Liés à différents facteurs dont, bien évidemment les mécanismes de planification discursive, ces derniers révèlent du même coup la mise en place progressive de la scène verbale. Il nous faudra donc distinguer ce qui relève de la relation de dominance, (dans : *j'ai parlé de la misogynie*), on a une relation de rectio explicitement entre le verbe et le complément, des constructions qui s'inscrivent plus ou moins explicitement dans des relations de parataxe (dans *c'est l'homme qui s'occupe de cette collection, qui publie souvent des dictionnaires*, les deux relatives sont en relation parataxique). La tâche n'est pas forcément aisée car des contextes d'entre deux peuvent s'imposer. Ainsi, dans *je suis né à Alger en 1943*, le deuxième circonstant est à la fois dominé par le verbe et en relation de parataxe avec le localisateur spatial.

2.3. Correspondance : unités pragmatiques-balises textuelles

La définition complète de nos unités pragmatiques et la pose des balises textuelles associées dérivent ainsi de l'organisation tridimensionnelle du message : (i) informationnelle, (ii) discursive et (iii) topologique⁴.

Concernant le premier niveau d'organisation, si nous partons des oppositions fondamentales posées par Lambrecht, nous nous éloignons ensuite de son cadre théorique, cela pour deux raisons : (i) nous opposons un thème neutre à un thème perceptivement marqué, car faisant l'objet d'une extraction⁵ ; (ii) les notions de **continuum focal** et de **dynamisme communicatif** (Firbas 1964), négligées par

⁴ Nous illustrons les différentes constructions autant que faire se peut par des exemples tirés du texte d'étude, enrichis, si besoin est, d'autres corpus de parole spontanée analysés.

⁵ Or, pour Lambrecht, un thème extrait ne peut pas être considéré comme marqué, car un élément marqué correspond forcément à une structure rare par rapport aux constructions fréquentes, ce qui n'est pas le cas de l'extraction thématique qui, selon l'auteur, est LA construction thématique par excellence en français parlé.

Lambrecht,⁶ sont ici essentielles pour traiter de la zone rhématique au sein de laquelle nous distinguons un rhème neutre et un focus. Soit les cinq étiquettes suivantes :

- Le thème neutre⁷ <THN> désigne un thème argumental qui, de fait, ne donne pas lieu à un détachement (*le dénigrement de la femme n'est pas une question de niveau social*).
- Le thème extrait à gauche <THG> représente une unité référentielle disloquée en position frontale, qu'elle soit liée ou non liée, *i.e.* co-référente ou non à un argument sujet ou objet dans la proposition⁸ (*Hollywood c'est quand-même l'image de Los Angeles, le chocolat j'adore*).
- La parenthèse <PAR> peut relever du parenthésage propositionnel (PAR_PROP), typique des constructions clivées où le focus est de nature argumentale (*c'est l'homme (í) qui publie souvent des dictionnaires*). Il peut s'agir également d'une unité référentielle détachée à droite, PAR_ARG, co-référente à l'argument sujet de la proposition (*elle n'est pas gentille, la maîtresse*).
- Dans le rhème neutre <RHN>, l'apport d'information est apporté par l'ensemble de la proposition (*ça a changé de forme*).
- Le rhème focalisé <FOCI> correspond prototypiquement à l'argument focus introduit dans une structure clivée. C'est lui qui est porteur du degré de dynamisme communicatif le plus haut.

La manipulation de la structure discursive, quant à elle, nous amène à envisager un jeu de quatre étiquettes :

- les deux premières, <TOPE> et <FOCE>, désignent une emphase prosodique (accentuation secondaire) susceptible d'être produite sur un topic (ou thème) et un focus, liée à l'intersubjectivité, (*tout le monde est*

⁶ La notion de dynamisme communicatif (DC) dérive d'une observation simple : les constituants autour desquels se construit l'information participent à des degrés divers au développement de la communication. En conséquence, plus un élément permet à l'information de se développer, plus son degré, dans l'échelle du dynamisme communicatif, sera élevé, d'où sa saillance perceptive. Si cette notion, relationnelle par nature, s'établit d'abord dans l'opposition dichotomique thème vs. rhème, elle concerne également les relations établies entre les segments de la zone rhématique. Lorsque l'on compare par exemple le DC du prédicat et du circonstant, nombreux sont les contextes où le verbe ne sert que de transition pour introduire l'information principale (*j'ai grandi dans l'insouciance*).

⁷ Désigné ainsi en référence à la phrase canonique [SN+V(+SN)], telle qu'elle est traitée par les grammairiens.

⁸ Ce type de construction est absent du corpus étudié.

mysogine). Le locuteur, souhaitant mettre en relief telle ou telle partie du discours, produit un accent d'insistance sur une unité indépendamment de sa charge informative et de sa saillance cognitive⁹. Ce type d'emphase est la plupart du temps prédictible dans le lexique, nous en voulons pour preuve le traitement prosodique des adverbes de modalité (*franchement, bien sûr*) et de point de vue (*pour moi*), celui des quantifieurs (*extrêmement, beaucoup, à quel point, tout le monde*) et de la négation (*non, pas du tout*), pour lesquels l'accent peut être codé dans le lexique.

- <TOPD> marque le verbe introducteur d'un discours direct rapporté (DRD), étiquette qui devra engendrer une saillance mélodique sur le verbe (*ils sont venus me direí , on m'œa dití*), suivie le cas échéant par une pause. Ces indices soulignent le rôle de cadrage thématique de la séquence antéposée au DRD et indiquent du même coup, le statut rhématique du DRD.
- <REP> indique un cadrage circonstanciel, préposé ou postposé au pivot verbal (*après un long travail de recherche, vous avez réuni plus de 2000 textes, ça a changé de forme aujourd'hui*). Pour éclairer ce point : si sous l'angle strictement syntaxique, un syntagme circonstanciel est souvent accessoire, *i.e.* non nécessaire pour que la phrase tienne debout syntaxiquement, il joue, sur le plan pragmatique, un rôle de repère fondamental¹⁰, qui permet au locuteur-auditeur de circuler dans le discours, et oriente ainsi la construction prédicative. De ce point de vue, le repère répond au principe de pertinence (Sperber & Wilson 1986) selon lequel un énoncé est d'autant plus pertinent que le coût cognitif pour le comprendre est faible : sa présence dans le discours, effectivement, permet de réduire ce coût.
- L'étiquette < SAM> est utilisée pour repérer une unité référentielle active dans la mémoire à court terme du sujet (degré de saillance mémorielle élevé), car déjà présente dans le cotexte gauche (*c'œst l'homme qui publie souvent des dictionnaires ó dictionnaire de la bêtise, dictionnaire des erreurs politiques, etc.*).

Enfin, trois balises topologiques sont manipulées :

⁹ On désigne par saillance cognitive d'une unité, son degré d'activation dans la représentation mentale d'un sujet.

¹⁰ Au sens de la Théorie des Opérations Enonciatives de Culioli (1990).

- <SAT> exprime la saturation fonctionnelle d'une tête quelle qu'en soit la nature (*je suis né à Alger en 1943, j'étais d'une famille très unie, très aisée*).
- <TAX> est associée à chaque unité intégrée dans une liste parataxique (*c'est très mauvais pour lui, pour sa virilité et pour son image*).
- L'unité qui ponctue la liste parataxique est dotée du marqueur final de liste : <FTAX> (*dictionnaire de la bêtise, dictionnaire des erreurs politiques, etcetera*).

Soit la séquence suivante : *c'est l'homme qui s'occupe de cette collection, qui publie souvent des dictionnaires*, le tableau ci-dessous nous permet d'illustrer notre utilisation des étiquettes telles qu'elles dérivent de nos trois niveaux de structuration du message : informatif, discursif et topologique.

Tableau 1 : Exemple de balisage pragmatique

Etiquette + occurrence	Fonction pragmatique
<FOCI> <i>C'est l'homme</i>	Argument focus fourni à une proposition ouverte présupposée : <i>quelqu'un s'occupe de cette collection</i> et réponse à la question partielle : <i>qui s'occupe de cette collection ?</i>
<PAR> <i>qui s'occupe de cette collection</i>	Rappel de la proposition ouverte présupposée : <i>quelqu'un s'occupe de cette collection</i>
<PAR> <TAX> <i>qui publie <FOCE> souvent/ des dictionnaires</i>	Spécification de la présupposition : <i>quelqu'un publie des dictionnaires</i>

où :

- Un focus informatif est prévu sur *c'est l'homme* qui a pour fonction d'instancier l'argument focus de la proposition ouverte présupposée *quelqu'un s'occupe de cette collection*,
- Les relatives ont toutes deux le statut de parenthèse, dont la présence ne sert qu'à consolider l'accord intersubjectif : il s'agit d'éléments thématiques pour lesquels le locuteur présuppose au début de son intervention qu'ils n'ont nul besoin de saillance parce qu'ils constituent déjà le thème de l'énoncé (*i.e.* l'élément présupposé).
- La présence simultanée de trois balises dans *qui publie souvent des dictionnaires* (parenthèse, liste parataxique, emphase) aura des répercussions à la fois sur l'ensemble de la chaîne syllabique pour la génération des accents (<PAR>) et sur des points de localisations précis, affectant à la fois le contour terminal (<TAX>) et une proéminence accentuelle intermédiaire (<FOCE>).

2.4. Du côté de la prosodie : de l'emboîtement à la rupture

Afin de répondre à notre question initiale concernant la nature des indices prosodiques qui permettent d'identifier des contrastes syntagmatiques et donc des îlots de saillance perceptive, regardons d'abord comment se construit notre structure intonative. En premier lieu, le continuum prosodique fourni en entrée est préalablement segmenté en groupes intonatifs successifs (GI) sur la base du repérage de prééminences accentuelles terminales (identification de variations mélodiques et ou de durée significatives sur la dernière syllabe pleine des mots). Sous l'angle phonologique ensuite, chaque configuration prosodique terminale est codée par une matrice de traits binaires où les informations sont fondées sur un principe de valeur relative. Selon ce principe, la valeur fonctionnelle des unités se définit d'abord à partir du repérage de positions paradigmatiques¹¹, ensuite en tenant compte de l'enchaînement syntagmatique des contours et de leur contexte immédiat. Autrement dit, la valeur d'un contour dépendra non seulement de la valeur des contours avec lesquels il pourrait commuter dans un contexte terminal, mais également de la configuration du contour terminal de groupe qui le précède et enfin de son entourage proche. Ainsi, la présence éventuelle d'une pause subséquente à un contour peut modifier la valeur de ce dernier et prendre elle-même différentes valeurs relativement à son environnement. L'interprétation fonctionnelle que l'on peut en faire dépend, en effet, de la présence éventuelle d'un euh d'hésitation antéposé ou postposé. Lorsque ce dernier est postposé, le profil montant ou descendant du contour intonatif qui le porte amène à attribuer deux valeurs distinctes à la pause (conclusive dans le premier cas, elle assume une fonction continuative dans le second). Quatre types d'informations sont donc codés : le profil d'un contour terminal de groupe intonatif, le niveau fréquentiel qu'il traverse, son caractère allongé ou non et son contexte droit (présence ou non d'une pause et/ou d'un euh d'hésitation).

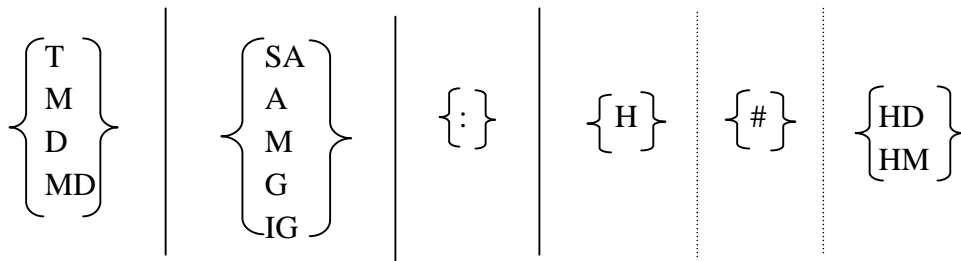


Figure 3 : matrice de traits associée aux contours terminaux

¹¹ Une unité ne prend sa valeur que relativement à d'autres qui sont substituables avec elle et qui forment son paradigme de définition.

où : de gauche à droite, sont indiqués : (i) les caractéristiques prosodiques d'un contour relatives à sa direction (montant, descendant, dynamique ou statique $\mp\emptyset$), le niveau traversé par un contour (de suraigu à infragave) et son caractère allongé ou non ; (ii) l'environnement droit du contour, le cas échéant : euh préposé à une pause (Hésitation), pause (silence ou inspiration), euh postposé porteur d'un contour descendant ou montant.

En somme, la matrice est utilisée pour déterminer les relations que deux groupes intonatifs contigus dans la chaîne sont susceptibles de contracter. Ce deuxième point, nous amène à construire une hiérarchie intonative à trois niveaux (voir Lacheret 2002b pour un exposé détaillé) qui s'exprime par quatre types de relation (figure 4) :

- Une relation de **liage direct** (GLD), au sein de laquelle plusieurs constituants syntaxiques fusionnent dans un seul groupe intonatif (*tout ceci est rassemblé dans un livre*).
- Dans la **relation de liage indirect** (GLI), un constituant syntaxique correspond d'abord à un et un seul groupe intonatif pour fusionner ensuite avec le GI contigu dans la chaîne (*il faudrait faire*) (*un dictionnaire de la misogynie*). L'emboîtement intonatif ici produit dérive d'un principe générique de dominance intonative qui donne lieu au regroupement des groupes intonatifs en paquets (pour l'origine de cette approche, voir Mertens 1987). Selon ce principe, tout GI domine le groupe intonatif immédiatement contigu à gauche, s'il est marqué soit par une excursion fréquentielle terminale dans le niveau aigu supérieure d'au moins un ton, soit par une descente dans le niveau infra-grave, soit par l'application d'un contraste de pente (voir Martin 1981 pour le statut fonctionnel de l'inversion de pente), soit, enfin, par la mise en œuvre d'un principe de progression marqué par un allongement d'au moins 50% de sa syllabe terminale par rapport à la syllabe finale du groupe intonatif qui le précède.
- Une **relation de rupture de premier niveau** (GR1) au sein de la même période intonative. Cette dernière, identifiée sur les bases de trois indices acoustiques majeurs (réinitialisation mélodique, présence d'une pause non contiguë à un « euh » d'hésitation, seuil d'amplitude terminale) désigne la racine de la structure prosodique. Lorsque les groupes intonatifs sont en relation de rupture à l'intérieur d'une période donnée, la disjonction est souvent explicitement marquée par une pause (*la misogynie*)#(*est le pilier*). Quand ce n'est pas le cas, on pose comme hypothèse que la chaîne est sous-spécifiée.
- La **relation de rupture de deuxième niveau** (GR2) désigne des contextes où un constituant forme un groupe intonatif autonome rejeté dans une

période contiguë à gauche ou à droite de sa tête périodique (*ça a changé de forme*)#(*aujourd'hui*).

Soit notre exemple : *c'est l'homme qui s'occupe de cette collection* et ses quatre configurations potentielles :

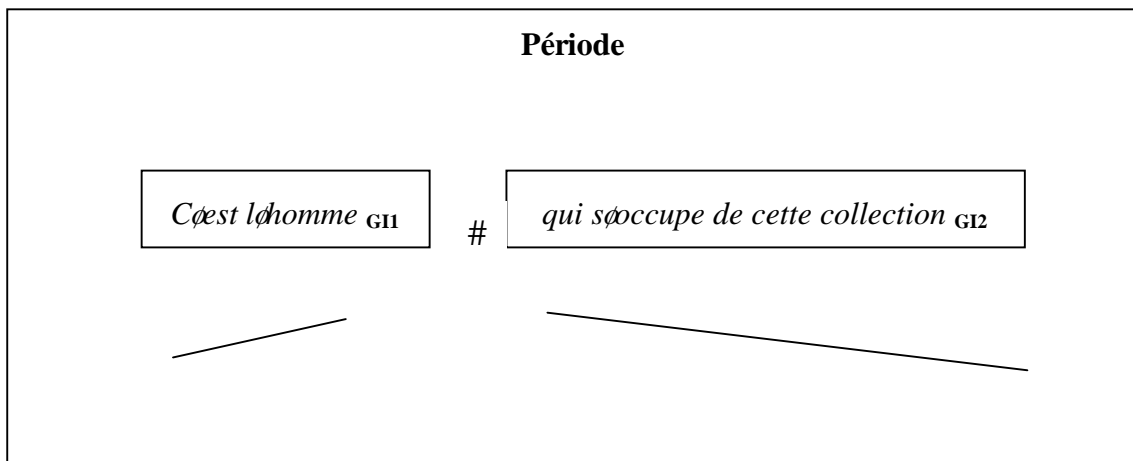
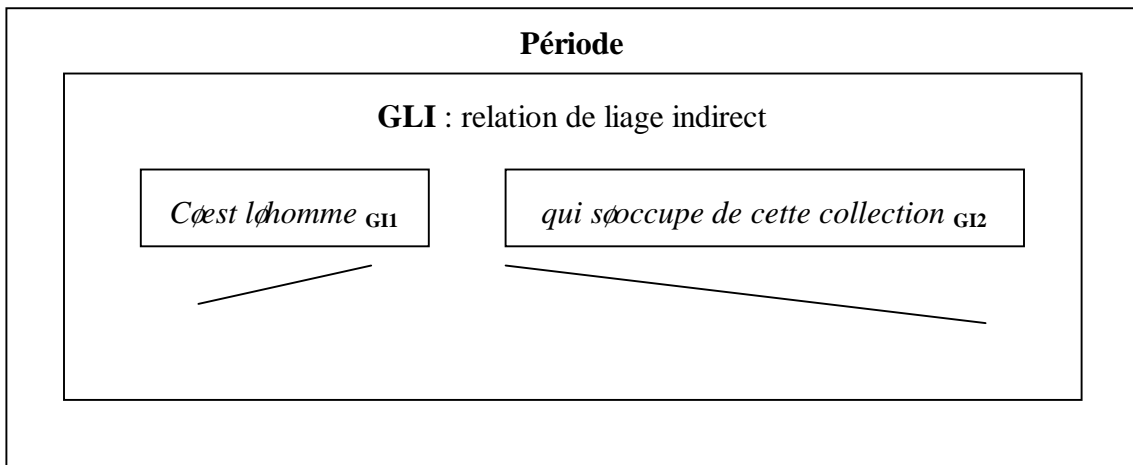
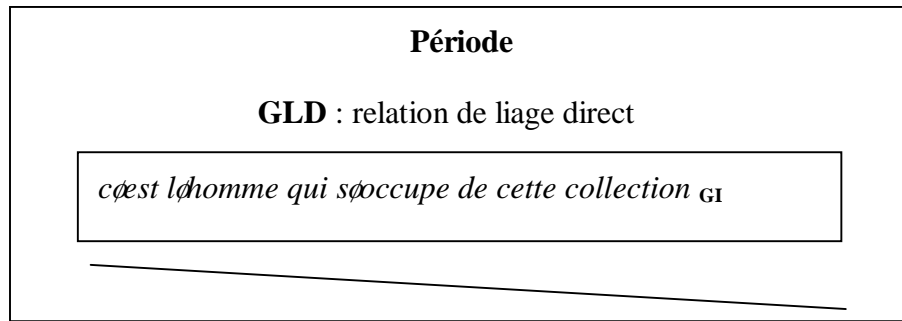




Figure 4 : les relations intonatives que peuvent contracter deux groupes intonatifs contigus, de l'emboîtement maximal à l'autonomie absolue

Nous proposons ensuite une **interprétation iconique** des objets prosodiques ainsi posés, qui s'exprime par la mise en œuvre d'un **principe de quantité** (Givon 1995), selon lequel, un contenu communicatif et conceptuel épais sollicite un codage linguistique dense. Lorsque ce dernier est faiblement exprimé, voire inexistant sous l'angle de la matérialisation syntaxique du message, la prosodie prend alors le relais ; autrement dit : ce que la syntaxe ne code pas, la prosodie le marque. Ce point nous conduit naturellement à réinvestir l'hypothèse de **l'effort articulatoire** débattue dans la littérature (Gussenhoven 2002) : il paraît logique de penser qu'un effort articulatoire est d'autant plus mobilisé qu'il faut rendre saillantes certaines unités du discours, c'est-à-dire produire des formes hyper-articulées qui se caractérisent par des traits prosodiques précis (dominance accentuelle, geste intonatif terminal ample et, par voie de conséquence, allongement significatif de la durée, excursion focale dans le niveau aigu ou suraigu, démarcation pausale). Mais, encore une fois, pour que l'intonation assume cette fonction de monstration, pour détecter des unités hyper-articulées, encore faut-il qu'elles s'opposent à des segments hypo-articulés (voir Di Cristo 1999 pour l'opposition entre formes hypo et formes hyper)¹². Pour ce faire, nous utilisons des indices à la fois paradigmatiques et syntagmatiques. Les premiers sont délivrés par le niveau de hauteur terminale de chaque unité, du suraigu à l'infra-grave, les seconds concernent l'enchaînement des groupes intonatifs (GLD, GLI, GIR1, GIR2). Plus précisément, les variations phonétiques de hauteur mélodique fusionnent en trois catégories phonologiques pertinentes : l'intonation focale pour les niveaux aigu et suraigu, l'intonation neutre pour les niveaux moyen et grave et l'intonation parenthétique pour le niveau infra-grave, soit les balises <FOC>, <NEU> et <PAR>.

¹² Ce point illustre les caractéristiques fondamentalement relatives des informations fournies par la matrice de traits (cf. *supra*, § 2.4.).

Les différentes configurations syntagmatiques ensuite (du GLD au GIR2) sont là pour forcer, voire réorienter, ces intonations par défaut. Ainsi, un constituant qui fait l'objet d'une autonomie accentuelle et qui est accompagné d'une démarcation prosodique est considéré comme un focus étroit (voir Ladd 1991), son degré de saillance est d'autant plus fort que la frontière prosodique qui le démarque est importante (voir la différence entre GR1 et GR2). Soit le continuum focal suivant :

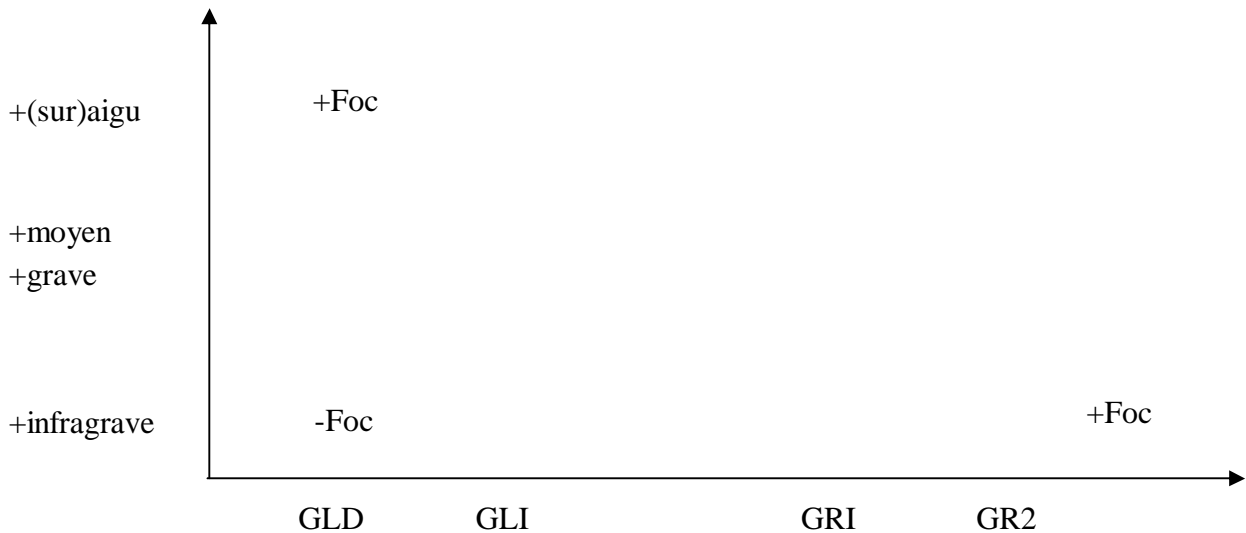


Figure 5 : le continuum focal et ses indices intonatifs

Enfin, en sachant que chaque contour peut être investi d'une fonction expressive, soulignons également l'inversion du patron logiquement attendu sur le thème (contour terminal descendant et non montant). En fait, tout comme l'extraction périodique d'un constituant présentée plus haut (construction GIR2), ce type de configuration ne peut pas se laisser réduire à une simple variante phonostylistique : l'utilisation d'un procédé inhabituel, non conforme aux règles de l'usage, permet de créer des points d'acrochage et de générer ainsi des centres d'attention perceptive. En somme, un thème dont le contour terminal s'organise autour des traits suivants [+suraigu, +GR2, +descendant] sera plus saillant que le même thème démarqué par un contour montant.

3. APPARIEMENT DES STRUCTURES

A partir de ces différents traitements, nous sommes en mesure de mettre en correspondance les marqueurs pragmatiques et prosodiques nécessaires pour passer de l'étape de l'analyse à celle de la génération des structures. Illustrons la méthode sur notre texte d'étude. En pratique, elle se décline en quatre points : (i) découpage du texte en groupes intonatifs, (ii) pause puis appariement des balises pragmatiques et

prosodiques (iii) sélection des patrons acoustiques les plus pertinents pour rendre compte de nos différents degrés de saillance perceptive lors du marquage du focus et de la parenthèse, (iv) intégration dans le système de synthèse vocale KALI développé au laboratoire CRISCO de l'université de Caen, dans le cadre d'un projet européen FEDER, et commercialisé depuis 1999. Nous ne nous attarderons pas sur les caractéristiques du traitement prosodique de KALI qui sont présentées de façon exhaustive dans Morel & Lacheret 2001 et rappelées dans ce numéro (voir l'article de Morel & Bänziger). Le point sur lequel nous souhaitons insister en revanche est le suivant : si le synthétiseur représente pour nous un outil pour mettre à l'épreuve notre modèle prosodique, en aucun cas notre objectif est de l'utiliser pour simuler la voix de B. Groult. Nous avons choisi Kali pour effectuer cette tâche car, certes son développement répondait d'abord à un objectif appliqué à commercialisation d'un système de synthèse pour mal-voyants et néanmoins, le système a également été conçu pour être un outil de simulation de nos traitements linguistiques. Dans ce cadre, deux versions ont été développées, la première, fermée, est destinée à la commercialisation, la seconde, ouverte, constitue une plate-forme de travail robuste pour entreprendre des tests psycho-acoustiques exhaustifs. L'accès direct et transparent aux trois paramètres prosodiques (f0, intensité, durée) permet en effet à l'expérimentateur de créer une combinatoire de paradigmes de tests extrêmement fine et approfondie : toutes les caractéristiques de la matrice de traits formulée lors de la modélisation phonologique (cf. *supra*, figure 3) sont contrôlables et susceptibles de manipulations¹³. De ce point de vue, la synthèse de la parole s'avère être un merveilleux outil de recherche, et ce, à double titre : (i) elle est utilisée pour mettre à l'épreuve les principes qui émergent de l'analyse en testant leur pertinence perceptive, donc cognitive, (ii) elle peut conduire à la remise en question d'évidences pourtant considérées comme acquises dans la littérature (voir en *infra*, § 4. la discussion autour du principe d'isochronie) et, à l'inverse, valider des hypothèses nouvelles, fécondes pour la génération des structures, ouvrant ainsi de nouvelles pistes de travail.

La méthode retenue pour conduire ce travail est simple ; en pratique, il s'agit de comparer deux types de signaux : (i) la voix synthétique telle qu'elle est générée automatiquement, (ii) la voix synthétique modifiée à l'issue du balisage manuel provenant du traitement pragmatique.

¹³ Nous sommes extrêmement reconnaissante envers Michel Morel qui a fait ici un travail colossal et extrêmement précieux.

3.1. De la pragmatique à la prosodie : quelles balises

Parmi les étiquettes pragmatiques associées à une séquence discursive, il faut distinguer celles qui affectent le patron global de la chaîne, de celles qui portent uniquement sur des proéminences accentuelles locales. Ainsi, <PAR> entraîne automatiquement la désaccentuation de l'ensemble de la chaîne syllabique qu'elle introduit, tandis que <TAX> affecte le patron terminal d'une unité au sein d'une structure de liste donnée. D'autres étiquettes sont associées à des points de localisation intermédiaire, c'est le cas du marqueur d'emphase <FOCE> qui conduit à la génération d'un accent secondaire, c'est-à-dire un accent non terminal pour les polysyllabes (figure 6).

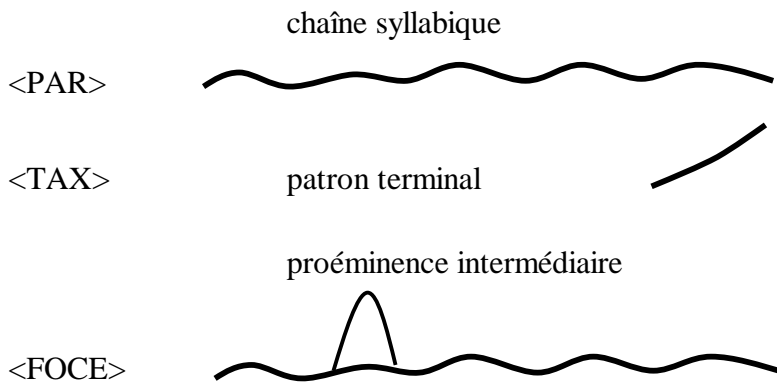


Figure 6 : zoom sur les étiquettes prosodiques

La comparaison des groupes intonatifs deux à deux (cf. *supra*, § 2.4.) nous conduit en définitive à faire émerger quelques principes de constructions prototypiques qui peuvent s'actualiser dans la substance de manière extrêmement riche et variable. Ces principes se résument comme suit (figure 7) :

- Un principe de symétrie maximale, somme toute relativement rare, dans lequel les gestes terminaux de groupes intonatifs se font écho de manière parfaitement symétrique, qu'il s'agisse de la direction des gestes, de leur ambitus ou de leur durée. Ce principe est mobilisé pour rendre compte des structures de liste, la symétrie étant rompue à l'issue de l'actualisation du dernier groupe, servant ainsi à indiquer la fin de la liste énumérative (*de Molière, à Baudelaire, à Maupassant, à Montherlant, à Dutour*).
- Un principe d'alternance qui correspond à un mode de production régulier et qui fait alterner figure et fond. Un tel principe valide l'hypothèse selon laquelle, les patrons prosodiques sont fondamentalement déterminés par un processus de génération asymétrique des structures, qui correspondrait sans

doute à une fonction cognitive élémentaire d'organisation de formes. Ce principe se décline à son tour en deux sous-principes :

- un principe d'alternance à dominance droite-gauche, le geste terminal du deuxième GI étant perceptivement plus saillant que le premier (*j'ai parlé de la misogynie*) (*dans ainsi soit-elle*).
- Un principe d'alternance à dominance gauche-droite, l'effet de saillance diminuant au fur et à mesure de la perception des cibles accentuelles terminales Il est notamment déterminé pour actualiser une parenthèse terminale (*c'est l'homme*) (*qui publie des dictionnaires*) (*dictionnaire de la bêtise*) *dictionnaire des erreurs politiques*).

La sélection de l'un de ce principes résulte de l'application d'un test récursif gauche-droite (comparaison du groupe 1 avec le groupe 2, du groupe 2 avec le groupe 3, etc.).

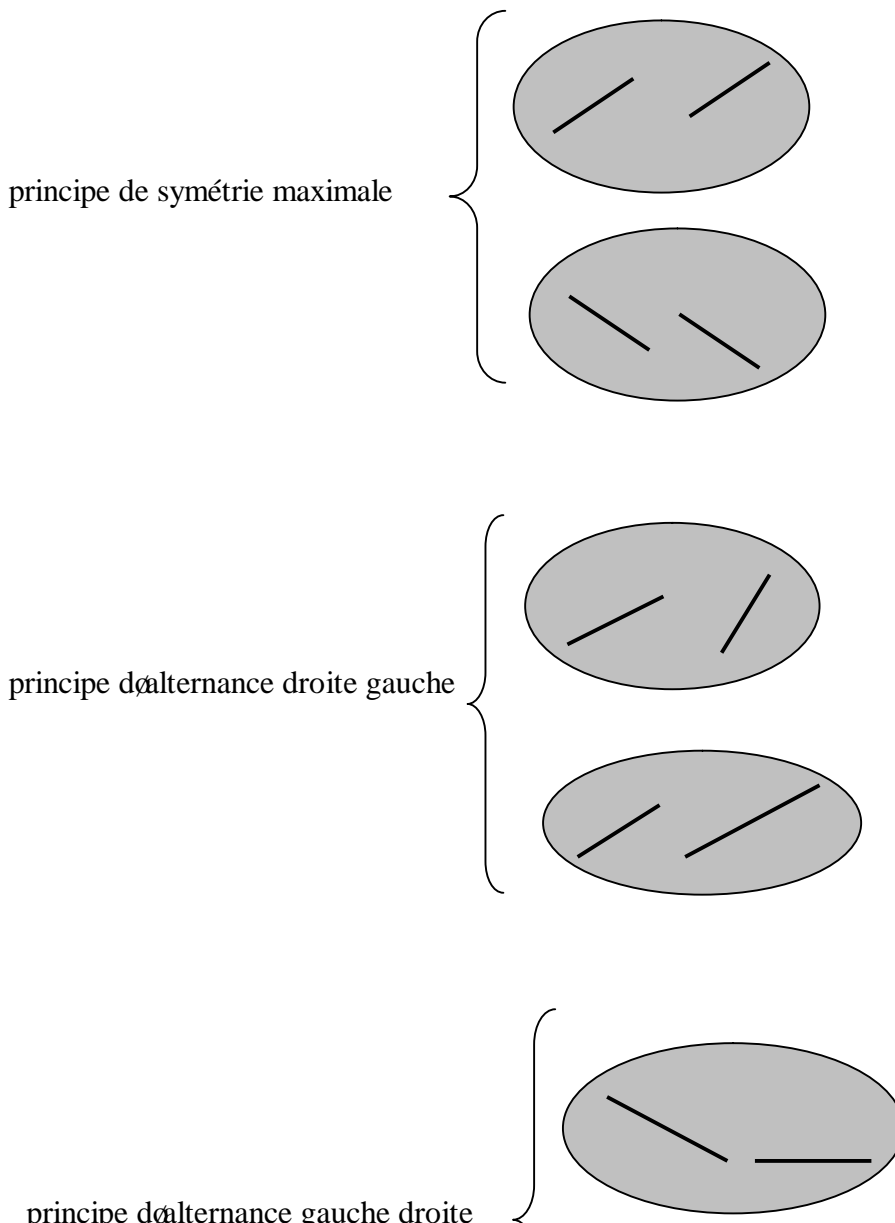


Figure 7 : schématisations intonatives, illustration des principes appliqués

Les problèmes qui se posent désormais sont de deux sortes, ils concernent d'une part la génération intrinsèque des contours, ensuite la modulation éventuelle de ces derniers en fonction de leurs contextes. Autrement dit :

- Pour une catégorie pragmatique donnée, comment générer un patron prosodique optimal ?
- Pour deux groupes intonatifs contigus, quels vont être les principes appliqués (alternance ou symétrie maximale), en vertu de la catégorie pragmatique à laquelle appartient chaque groupe ?
- Quels sont les corrélats acoustiques de ces patrons abstraits concernant la f_0 , l'intensité et les variations de durée (segments pleins et pauses) ?

En somme, notre tâche consiste à effectuer le meilleur appariement entre les balises pragmatiques posées, les balises prosodiques et les principes de construction intonatives.

Le codage acoustique fait intervenir sept champs (tableau 2), chaque champ pouvant être codé par cinq valeurs¹⁴ : (i) 0 indique que l'on ne modifie par la prosodie par défaut générée dans Kali, les paramètres F_0 , intensité, durée ne sont donc pas modifiés, (ii) les marqueurs + et ++ nous conduisent à augmenter de 50% (+) la prosodie par défaut, voire de 100% (++)¹⁵, (iii) l'opération inverse est indiquée par le marqueur ÷÷. Les modifications de direction de contour sont également indiquées. Ainsi, le marqueur ↯ signifie qu'un contour descendant, prévu dans le système de génération automatique, se substitue un contour montant.

¹⁴ Les paramètres pertinents pour nos manipulations sont la fréquence fondamentale et la durée, l'intensité, étant considérée pour l'heure comme une co-variable de la F_0 , n'est pas traitée indépendamment de cette dernière.

¹⁵ Ces expériences nous ont montré que pour obtenir un effet perceptivement significatif, la valeur de base devait être modifiée de façon notable, d'où l'utilisation fréquente du marqueur ÷÷÷

Tableau 2 : champs prosodiques retenus pour le traitement acoustique

1. Accent secondaire
2. Intonation parenthétique des syllabes intermédiaires
3. Direction d'un contour terminal
4. Ambitus (plein ou demi)
5. Allongement ou compression de la durée par défaut
6. Augmentation ou diminution du temps de pause
7. Dislocation périodique possible ¹⁶

3.2. Illustration sur un extrait du corpus

Pour illustrer ce traitement, reprenons notre exemple introduit au paragraphe 2.3. (cf. *supra*, tableau 1). Tout d'abord, les étiquettes pragmatiques sont réécrites en étiquettes prosodiques, ici :

- {<FOCI>, <FOCE>} → <FOC>
- <PAR> → <PAR>

Ensuite, les enchaînements pragmatiques auxquels correspondent deux groupes intonatifs contigus conditionnent la sélection d'un principe de construction intonative particulier (SYM, PAD ou PAG) de la façon suivante :

- <FOC> + <PAR> → <PAG> : où la saillance d'un patron terminal d'un focus intonatif dominera par rapport à la saillance perceptive faible de la parenthèse qui le suit.
- <TAX> → <SYM> : où les constructions prosodiques des constituants intonatifs, s'inscrivant dans une structure de liste, sont symétriques.

¹⁶ Afin d'éviter la redondance des marques, lorsque le champ <dislocation périodique> est rempli, le champ <pause> ne l'est pas, puisque le premier entraîne automatiquement l'insertion d'une pause.

Tableau 3 : marqueurs prosodiques et corrélats pragmatiques

Etiquettes prosodiques + occurrence	Codage acoustico-phonétique
<FOC> <i>Cøest løhomme</i>	[0 ;0 ;{M,D} ;++, (+) ;0 ;(+)]
<PAR> <i>qui søoccupe de cette collection</i>	[0 ;- ;{M,D} ;0 ;(+) ;0 ;(+)]
FOC + PAR → PAG	[0 ;- ;M ;0 ;++ ;0 ;+] [0 ;- ;D ;0 ;+ ;0 ;0]
<PAR> <TAX> <i>qui publie <FOC> souvent/ des dictionnaires</i>	[+ ;- ;{M,D} ;0 ;(+) ;0 ;(+)]
TAX → SYM	[0 ;- ;D ;0 ;+ ;0 ;0] [+ ;- ;D ;0 ;+ ;0 ;+]

Avec la lecture suivante :

- *Cøest løhomme* : pas d'accentuation secondaire, pas de parenthèse, intonation du contour terminale montante ou descendante en variation libre, ambitus plein du contour, allongement vocalique terminal de 50% envisageable par rapport à l'allongement prévu par défaut ; le temps de pause reste inchangé ; la dislocation périodique possible permettrait ici de mettre en valeur de manière optimale le focus pragmatique *homme*.

- *Qui søoccupe de cette collection* : le statut de parenthèse ici déclenche une diminution des modulations de fréquence fondamentale générées par défaut (champ 2), pour le reste, les modifications prévues sont facultatives, donc entre parenthèses. On peut ainsi, si on le souhaite, augmenter l'allongement vocalique (champ 5) et introduire une dislocation périodique (champ 7).

- **Foc + PAR → PAG** : on aligne les patrons mélodiques des deux groupes intonatifs (*cøest løhomme*) + (*qui søoccupe de cette collection*) et on sélectionne les variantes pertinentes en vertu du principe appliqué. Ainsi, l'intonation focale du premier GI suivie d'une intonation parenthétique nous amènerons à sélectionner l'enchaînement intonatif terminal [Montant + Descendant]. Les variantes intonatives proposées au départ, qui ne s'imposent pas du point de vue des principes activés, ne sont finalement pas prises en compte ((+) → 0)).

- Le même processus est appliqué pour générer l'enchaînement intonatif correct des groupes (*qui søoccupe de cette collection*) suivi de (*qui publie souvent des dictionnaires*). Etant donné l'application du principe de symétrie maximale, les contours terminaux ne pourront être que montants ou descendants ; le contour descendant ayant été choisi pour le premier GI lors de la phase précédente de traitement, ce contour est également sélectionné pour le second. Un traitement identique est appliqué pour le calcul de la durée. Seul le dernier champ est différent : la fin de la structure de liste parenthétique est marquée par une dislocation périodique. Pour finir, précisons que l'intonation parenthétique prévue pour les deux relatives n'exclue des prééminences locales. Ainsi, *souvent* reçoit un accent secondaire sur sa première syllabe (<FOC> souvent/).

4. CONCLUSION

La problématique abordée dans le cadre de l'appel d'offre du Groupe plurifacultaire Prosodie de l'Université de Genève nous a amenée à nous confronter à de vraies questions, tant du côté de la génération que de l'analyse¹⁷.

Sous l'angle de la génération d'abord, c'est-à-dire de la pertinence de nos balises pour une génération automatique de la prosodie moins artificielle et moins décevante qu'elle ne l'est encore aujourd'hui. Il serait illusoire de conclure que les balises, posées manuellement dans ce travail, peuvent toutes donner lieu à l'implémentation. Une automatiser complète du traitement proposé repose sur des analyses syntaxiques robustes qui ne sont pas disponibles pour l'heure. Par analyse robuste, nous entendons des systèmes qui n'achoppent pas pour calculer correctement les relations de dépendance à distance (*c'est l'homme qui s'occupe de cette collection qui publie souvent des dictionnaires*). Ce point peut sans doute être partiellement résolu par un traitement exhaustif de la ponctuation. En revanche, les structures de liste parataxique posent encore de réels problèmes aux analyseurs. Ces dernières ne pourront être correctement traitées qu'en incluant des informations sémantiques. Autre aspect à défricher : quel choix retenir en définitive, dans les contextes de variations ? Deux solutions sont envisageables. La première, sans doute directement opératoire, conduirait à activer le patron intonatif qui modifie le moins la prosodie par défaut. Une autre solution, plus subtile, mais à quel prix ?, consiste à intégrer autant que faire se peut, l'ensemble des indices énonciatifs, à les formaliser par des règles ou des marqueurs directement intégrés dans un lexique. Si cette approche n'est bien évidemment pas applicable à toutes les unités du discours, elle permettrait notamment un meilleur traitement des adverbes souvent fondamentalement liés à des enjeux co-énonciatifs¹⁸. Ainsi, pour le marqueur de discordance de point de vue¹⁹ *non pas du tout*, nous pourrions substituer au contour montant, généré automatiquement, un contour descendant dans le niveau infra-grave, marquant ainsi une rupture dialogale. Bref, comme toujours on se trouve pris entre deux impératifs *a priori* contradictoires : une recherche de solution opérationnelle au moindre coût et un souci de pertinence théorique.

¹⁷ Nous ne pouvons que louer l'initiative de nos collègues genevois, qui se faisait attendre depuis de nombreuses années au sein de la francophonie. Sans aucun doute, avons-nous tous, théoriciens, modélisateurs et spécialistes de synthèse de la parole grand intérêt à systématiser ce genre de rencontre, gage de recherches fécondes.

¹⁸ L'importance intonative de ces derniers a d'ailleurs été intuitivement reconnue et traitée par d'autres auteurs (Martin 1995, Mertens 2002).

¹⁹ Nous empruntons ce terme à Morel et Danon-Boileau (1998).

Concernant la validation des modèles et des théories intonatives ensuite, les manipulations psycho-acoustiques, réalisées à l'occasion de ce travail, nous ont amené à faire le point sur l'ensemble de principes de construction accentuelle et intonative susceptibles d'être définis, et à en revisiter certains, pourtant considérés comme acquis dans la littérature, alors que d'autres, plus marginaux, apparaissent pourtant comme essentiels. Prenons l'exemple des principes d'équilibre syllabique et de progression *a priori* antagonistes. Selon le premier, (1) à l'intérieur d'un groupe intonatif donné, les syllabes inaccentuées ont une durée moyenne identique, (2) au sein d'une séquence prosodique, chaque groupe intonatif possède, autant que faire se peut le même nombre de syllabes. Si le premier est posé comme un principe incontournable par les phonéticiens, il est pourtant contradictoire avec la remise en question de l'isochronie syllabique en français. Selon cette dernière, les syllabes inaccentuées d'un groupe seraient en fait la culmination d'un mouvement auquel participent tous les éléments de ce groupe, la durée des syllabes inaccentuées tendant à progresser jusqu'à la syllabe accentuée dans un mouvement général de ralentissement progressif (Padeloup 1990, Lacheret & Beaugendre 1999). Le principe d'équilibre s'oppose donc à un principe de progression, qui se vérifie également concernant le nombre de syllabes de deux groupes intonatifs contigus. Sur ce point, on a pu noter en effet que : au sein d'une même séquence intonative, le nombre de syllabes de chaque GI peut augmenter de façon constante (*des pâtes*) (*des pâtes*) (*oui mais des Panzani*). Ainsi, les auteurs qui mettent en exergue le principe d'équilibre rythmique sont les mêmes à envisager le phénomène de progression comme un vecteur essentiel des constructions rythmiques en français (Padeloup 1990, Delais 1995). Pour résoudre cette contradiction, Delais (1995) envisage de modéliser ces principes dans le cadre de la théorie de l'optimalité. Pourtant, dans la hiérarchie définie par l'auteur, les deux contraintes sont en variation libre. Il nous semble là qu'il s'agit d'une manière relativement rapide de régler le problème. Certes, sur le plan rythmique, les deux principes sont en variation libre, mais du point de vue des contraintes pragmatiques, il apparaît évident que le principe de progression domine et qu'il dépend d'une contrainte beaucoup plus fondamentale encore, sans doute universelle (van Valin & Lapolla 1997), qui consiste toujours à commencer par le groupe le moins informatif, donc *a priori* le plus court, pour finir par le groupe le plus informatif, c'est-à-dire le plus long²⁰. Pour revenir à la remise en question de l'isochronie syllabique en français, totalement validée ici par nos manipulations acoustiques, elle soulève une fois de plus le caractère déterminant de la durée dans l'actualisation phonétique de la structure intonative, tant pour le marquage de l'accent secondaire, de nature focalisante (*faites votre enquête, interrogez les*

²⁰ Nous en voulons pour preuve, dans des langues comme le français du moins, la position rhématique droite prototypique, la position initiale étant, par essence, le site thématique privilégié.

jeunes)²¹ que pour celui de l'accent primaire. En somme, il reste beaucoup à faire pour disposer de modèles robustes traitant de l'épineux problème de l'élasticité temporelle en français (Zellner 2002), et de la façon dont elle affecte la focalisation accentuelle. Enfin, de manière tout aussi fondamentale, ces travaux ont montré l'impact de l'effort articulatoire mobilisé sur une cible syllabique, en vue de produire un accent de focalisation, sur les syllabes adjacentes (voir aussi Brichet & Aubergé 2002 pour des conclusions similaires). La production d'un accent, en effet, entraîne irrémédiablement une déformation, parfois significative, des syllabes environnantes. Ce phénomène, qui s'explique très bien en terme de co-articulation, illustre du même coup la pertinence cognitive des modèles globaux, raisonnant en termes de contours prototypiques (Aubergé 1991), et sans aucun doute leur efficacité pratique, au détriment des approches locales qui s'en tiennent à la manipulation de points cibles. C'est dans cette voie que nous souhaitons poursuivre nos travaux pour la génération automatique des structures prosodiques au laboratoire CRISCO de l'université de Caen.

Anne LACHERET-DUJOUR
 Université de Caen, CRISCO
 Esplanade de la Paix 14032 Caen cedex
 France
 Anne.lacheret@crisco.unicaen.fr

RÉFÉRENCES BIBLIOGRAPHIQUES

- ANDERSEN Leth Hanne & Henning Nølke (dir.). 2002. « Macro-syntaxe et macro-sémantique », Bern, Peter-Lang.
- AUBERGE, Véronique. 1991. *La synthèse de la parole : des règles aux lexiques*, Thèse de Doctorat, Université de Grenoble.
- AUBERGE, Véronique, Anne LACHERET, Hélène LOEVENBRUCK (dir, sous presse). 2002. *Actes des Journées d'Étude sur la Prosodie*, Université de Grenoble.
- AUSHLIN, Antoine & A FERRARI. 1994. « Structuration prosodique, syntaxe, discours : évidences et problèmes », in *Cahiers de Linguistique française*, 15, Genève, 187-216.

²¹ Des allongements peuvent également être générés en position non finale (*c'est vraiment super*).

- BERRENDONNER, Alain. 1997. « Schématisation et topographie imaginaire du discours », in D. Miéville & A. Berrendonner (dir.), Bern, Peter Lang, 219-238.
- BERRENDONNER, Alain. 2002. « Morpho-syntaxe, pragma-syntaxe, et ambivalences sémantiques », in H.L. Andersen & H. Nølke (dir.), 23-42.
- BLANCHE-BENVENISTE, Claire (dir.). 1991. *Le français parlé*, Paris, éditions du CNRS.
- BRICHET, Cécile & Véronique Aubergé (sous presse). 2002. « La focalisation en français : morphologie de la prosodie », in V. Aubergé, A. Lacheret & H. Loevenbruck (dir.).
- CHAFE, Wallace W. (1987) : « Cognitive Constraints on Information Flow », *Coherence and Grounding in Discourse*, R. Tomlin (dir.), Amsterdam, Benjamins, 21-51.
- CULIOLI, Antoine. 1990. *Pour une linguistique de l'énonciation*, Paris, Ophrys.
- D'ALESSANDRO, Christophe & Evelyne TZOUKERMANN (dir.). 2001. *Synthèse de la parole à partir du texte*, Revue TAL, 42, Paris, Hermes.
- DELAIS, Elisabeth. 1995. *Pour une approche probabiliste de la structure prosodique, étude de l'organisation rythmique de la phrases française*, Thèse de Doctorat, Université de Toulouse.
- DI CRISTO, Albert. 1999 : « La cadre accentuel du français contemporain : essai de modélisation », *Langues*, vol2, n°3 et 4, Aupelf-Uref (dir.), 184-205, 258-269.
- FIRBAS, Jan. 1964). « On defining the Theme in Functional Sentence Perspective », *Travaux linguistiques de Prague*, 2, 267-280.
- FUCHS, Catherine & Stéphane Robert (dir.). 1997. *Diversité des langues et représentations cognitives*, Paris, Ophrys.
- GIVON, Talmy. 1994. « Isomorphism in the grammatical code, cognitive and biological considerations », in R. Simone (dir.), 47-76.

GRIZE, Jean-Blaise. 1990. *Logique et langage*, Paris, Ophrys.

GUSSENHOVEN, Carlos. 2002. « Intonation and Interpretation : Phonetic and Phonology », *Speech Prosody*, Aix-en-Provence, D. Hirst (dir.), 47-57.

HIRST, Daniel. & Albert Di Cristo (dir.). 1998. *Intonation Systems. A Survey of Twenty Languages*, Cambridge University Press.

LACHERET-DUJOUR, Anne & Frédéric Beaugendre. 1999. *La prosodie du français*, éditions du CNRS.

LACHERET-DUJOUR, Anne. 2002a. « The Intonational Marking of Topical Saliency in Spontaneous Speech : Evidence from spoken French », *Speech prosody 2002*, D. Hirst (dir.), 447-450.

LACHERET-DUJOUR, Anne. 2002b. *Modélisation prosodique du français parlé : analyse de la substance, représentation formelle, interprétation symbolique*, Travaux de Habilitation à diriger des recherches, Paris X, Nanterre.

LACHERET-DUJOUR, Anne. (sous presse). 2002.c. « Modéliser l'intonation d'une langue. Où commence et où s'arrête l'autonomie du modèle ? L'exemple du français parlé », *Journées Prosodie 2001*, V. Aubergé, A. Lacheret, H. Loevenbruck (dir.), Grenoble.

LACHERET-DUJOUR, Anne & Bernard VICTORRI (sous presse). 2002. « la période intonative comme unité d'analyse pour l'étude du français parlé : modélisation prosodique et enjeux linguistiques », in *Verbum*, M. Charolles (dir.), Nancy.

LAKS, Bernard. 1997. *Phonologie accentuelle et Métrique, autosegmentalité et constituance* et Paris, éditions du CNRS.

LAMBRECHT, Knud. 1994, *Information Structure and Sentence Form*, Cambridge University Press.

MARTIN, Philippe. 1981. « Pour une théorie de l'intonation : l'intonation est-elle une structure congruente à la syntaxe ? » in M. Rossi & al. (dir.), 234-271.

- MARTIN, Philippe. 1985. « Interaction prosodie-syntaxe en français : cas des adverbes en énoncé », in Ch. Sorin & al. (dir.), 171-182.
- MERTENS, Piet. 1987. *L'intonation du français : de la description linguistique à la reconnaissance automatique*, Thèse de Doctorat, Université de Louvain.
- MERTENS, Piet, Jean-Philippe Goldman, Eric Wehrli, Arnaud Gadinat. 2001. « La synthèse de l'intonation à partir de structures syntaxiques riches », in revue TAL, Ch. D'Alessandro & E. Tzoukermann (dir.), 145-193.
- MERTENS, Piet. (soumis pour publication). 2002. « A classification of French Adverbs based on Distributional, Syntactic and Prosodic Criteria ».
- MIEVILLE, Denis & Alain BERRENDONNER (dir). 1997. *Logique, discours et pensée*, Berne, Peter Lang.
- MOREL, Marie Annick & Laurent DANON-BOILEAU. (1998). *Grammaire de l'intonation à l'exemple du français*, Paris, Ophrys.
- MOREL, Michel & Anne LACHERET-DUJOUR. 2001. « Kali : synthèse vocale à partir du texte. De la conception à la mise en œuvre », in revue TAL, Ch. D'Alessandro & E. Tzoukermann (dir.), 193-222.
- PASDELOUP, Valérie. 1990. *Modèle de règles rythmiques du français appliqué à la synthèse de la parole*, Thèse de Doctorat, Université d'Aix-en-Provence.
- PREVOST, Sophie. 2001. *La postposition du sujet en français aux XVe et XVIe siècles, analyse sémantico-pragmatique*, Paris, CNRS Editions.
- ROBERT, Stéphane. 1997. « Variation des représentations linguistiques : des unités à l'énoncé », in C. Fuchs & S. Robert (dir.), 25-37.
- ROSSI, Mario. & al. (dir.). 1981. *L'intonation : de l'acoustique à la sémantique*, Paris, Klincksieck.
- ROSSI, Mario. 1981a. « Vers une théorie de l'intonation », in M. Rossi & al. (dir.), 179-183.

- ROSSI, Mario. 1981b « Intonation, énonciation, syntaxe », in M. Rossi & al. (dir.), 184-233.
- ROSSI, Mario. 1981c. « L'intonation n'est pas congruente à la syntaxe : une explication », in M. Rossi & al. (dir.), 290-296.
- ROSSI, Mario. 1999. *L'intonation, le système du français ó description et modélisation*, Paris, Ophrys.
- SORIN, Christel & al. (dir.). 1995. *Levels in speech communication, Relations and interactions*, Amsterdam, Elsevier.
- SELKIRK, Elisabeth. 2001. « The interaction of constraints on prosodic phrasing, in Prosody : Theory and Experiment », M. Horne (dir.), Dordrecht : Kluwer Publishing.
- SIMONE, Rapahel. (dir.). 1994. *Iconicity in Language*, Amsterdam Benjamin.
- SPERBER, Dan. & Deirdre WILSON. 1986. *La pertinence*, Paris, Minuit.
- TALMY, Leonard. 2000. *Toward a Cognitive Semantic*, vol 1, Cambridge, MIT Press.
- Van VALIN, Robert & Randy J.R. LAPOLLA. 1997. *Syntax : structure, meaning and function*, Cambridge, Cambridge University Press.
- VICTORRI, Bernard & Catherine FUCHS. 1996. *La polysémie. Construction dynamique du sens*, Paris, Hermès.
- VICTORRI, Bernard. 1999. « Le sens grammatical », in *Langages* n°136, 85-105.
- ZELLNER, Brigitte (sous rpesse). 2002. « Les enjeux de la simulation scientifique, l'exemple du rythme de la parole », in V. Aubergé, A. Lacheret-Dujour, H. Loevenbruck (dir.).